DOCUMENT RESUME

ED 104 898 TM 004 340

TITLE	Individual Pupil Report: Explanatory Materials. The Second Report of the 1971-72 Michigan Educational Assessment Program.
INSTITUTION	Michigan State Dept. of Education, Lansing. Research, Evaluation, and Assessment Services.
PUB DATE	Apr 72
NOTE	35p.; For related documents, see TM 004 339-342 and 555-556
EDRS PRICE	MF-$0.76 HC-$1.95 PLUS POSTAGE
DESCRIPTORS	Achievement Tests; *Basic Skills; Comparative Statistics; *Educational Assessment; Elementary Education; Elementary School Students; Raw Scores; Scores; *State Programs; Statistical Analysis; *Testing; *Test Interpretation; Test Reliability; Test Validity
IDENTIFIERS	*Michigan Educational Assessment Program

ABSTRACT

This explanatory booklet and the materials that accompany it have two major purposes. The first purpose is to provide local school officials with information regaing the performance on basic skills achievement of each student who took the 1971-72 Michigan Educational Assessment Battery. The second purpose is to provide local officials with information that will assist them in understanding and interpreting their students' scores. This booklet has four major sections. The first section describes the content of each test--word relationships, reading, mechanics of written English, and mathematics--in the educational assessment battery. The second section describes cautions that must be exercised in the interpretation of individual pupil scores from the program. The third section explains how to interpret the materials that accompany this booklet. Explanations are provided for pupils' scores. The fourth section defines statistical terms used in the educational assessment program and provides technical information regarding the educational assessment battery. (Author/DEP)

ED104893

INDIVIDUAL PUPIL REPORT: EXPLANATORY MATERIALS

The Second Report of the 1971-72
Michigan Educational Assessment Program

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TM 004 340

Prepared by Research, Evaluation and Assessment Services

Michigan Department of Education

April, 1972

2/3

State Board of Education

	Term Expires
Edwin L. Novak, O.D., *President* Flint	Jan. 1, 1973
Michael J. Deeb, *Vice President* Detroit	Jan. 1, 1977
Dr. Gorton Riethmiller, *Secretary* Chelsea	Jan. 1, 1975
Thomas J. Brennan, *Treasurer* Dearborn	Jan. 1, 1979
Marilyn Jean Kelly Detroit	Jan. 1, 1977
Annetta Miller Huntington Woods	Jan. 1, 1979
Dr. Charles E. Morton Detroit	Jan. 1, 1973
James F. O'Neil Livonia	Jan. 1, 1975
Dr. John W. Porter, *Superintendent of Public Instruction*, Chairman, Ex-Officio	
William G. Milliken, *Governor* Member, Ex-Officio	

FOREWORD

The Michigan Educational Assessment Program was initiated by the State Board of Education, supported by the Governor, and funded by the legislature initially through enactment of Act 307 of the Public Acts of 1969, and subsequently under Act 38 of the Public Acts of 1970. This report, the second in the 1971-72 series, provides information which will assist school district staff in the interpretation of the educational assessment results for individual pupils.

The State Board of Education has adopted a six-step process as a guide or model for improving Michigan education. The six steps are: the identification of common goals, the development of performance objectives, the assessment of educational needs, the analysis of delivery systems, the evaluation and testing of these systems or programs, and recommendations for educational improvement. This report presents information for the third step--the assessment of educational needs. Educational assessment provides general information on student needs which, along with other information gathered by local educators, will assist in identifying areas of need on the part of local schools and pupils. Analysis of the systems for delivering educational services and the specific evaluations of the areas so identified may then be initiated by local school officials in order to determine the extent to which changes in curricula and resource allocations are justified. Thus, the educational assessment program can contribute to the improvement of educational programs for Michigan children and youth.

Thanks are due to a large number of individuals and groups for making the Michigan Educational Assessment Program a reality and for continuing to work with it in its third year, 1971-72: to the State Board of Education for initially proposing it and continuing to support it, to the Governor and legislature for actively supporting it, and to Michigan educators for assisting with it. The program was designed and administered by the Research, Evaluation, and Assessment Services Unit, Michigan Department of Education, with the assistance of Educational Testing Service of Princeton, New Jersey, and the counsel of several ad hoc advisory groups.

This report was prepared by Dr. David Donovan, Mr. Robert Huyser, Dr. Philip Kearney, Mrs. June Olsen, and Dr. Daniel E. Schooley. Questions or requests for additional information relative to this report should be directed to the educational assessment staff.

John W. Porter
Superintendent of
Public Instruction

TABLE OF CONTENTS

SECTION		PAGE
	INTRODUCTION	1
I.	DESCRIPTION OF THE EDUCATIONAL ASSESSMENT MEASURES	2
	Word Relationships	2
	Reading	3
	Mechanics of Written English	3
	Mathematics	3
	Composite Achievement	4
II.	PRECAUTIONS IN THE INTERPRETATION OF PUPIL SCORES	5
	Measurement Error	5
	Content is General	5
	Results not Diagnostic	6
III.	INTERPRETATION OF PUPIL SCORES	7
	Description of Norming Group	7
	Definition of Terms	7
	Description of Pupil Roster and Pupil Labels	9
IV.	STATISTICAL CHARACTERISTICS OF THE 1971-72 MICHIGAN EDUCATIONAL ASSESSMENT BATTERY	12
	Definition of terms	12
	Statistical Characteristics of the Fourth Grade Basic Skills Battery	16
	Statistical Characteristics of the Seventh Grade Basic Skills Battery	19
	Summary	22
	APPENDIX	23
	Tables of Scores and Percents Below	24
	Means and Standard Deviations for the Norming Group	29

INTRODUCTION

This explanatory booklet and the materials that accompany it have two major purposes. This first purpose is to provide local school officials with information regarding the performance on basic skills achievement of each student who took the 1971-72 Michigan Educational Assessment Battery. The second purpose is to provide local officials with information that will assist them in understanding and interpreting their students' scores.

This booklet has four major sections. The first section describes the content of each test--word relationships, reading, mechanics of written English, and mathematics--in the educational assessment battery. The second section describes cautions that must be exercised in the interpretation of individual pupil scores from the program. The third section explains how to interpret the materials that accompany this booklet. Explanations are provided for pupils' scores. The fourth section defines statistical terms used in the educational assessment program and provides technical information regarding the educational assessment battery.

SECTION I

DESCRIPTION OF THE EDUCATIONAL ASSESSMENT MEASURES

The 1972 educational assessment battery was developed according to specifications developed jointly by Educational Testing Service and the Michigan Department of Education and reviewed and revised where necessary by Item Development Panels.[1] On the advice of these panels, specific changes were made in a number of the questions in this year's tests. Some of the changes suggested by panel members could not be implemented and still retain enough questions from the previous test to permit comparison of test results between years.

The 1971-72 educational assessment battery consisted of four different achievement measures--word relationships, reading, mechanics of written English, and mathematics. Each measure was a separate, timed test at the fourth and seventh grade. It is the purpose of this section to describe these tests.

Word Relationships

The fourth grade word relationships test contained 45 verbal analogy problems which were designed to measure students' knowledge of the meaning of words and the relationships between words and concepts. The seventh grade test contained 50 questions of the same type. The time allowed to work on this test was 20 minutes at both grades.

---

[1]Department members and an Educational Testing Service representative were also present at the panel meetings. The names and affiliations of these panel members are available upon request from the Bureau of Research, Evaluation and Assessment.

Reading

The reading test contained 50 questions which assessed paragraph comprehension, ability to understand words from the context in which they are encountered, and ability to identify the correct synonym for a word. Students at both grade levels were allowed 35 minutes to work on this test.

Mechanics of Written English

The mechanics of written English test consisted of four parts for fourth graders, and three for seventh graders, each separately timed. In part A, spelling, students were to identify misspelled words. The fourth grade test presented 15 items to be completed in 5 minutes; the seventh grade test presented 30 items to be completed in 6 minutes. Part B, effectiveness of written expression required students to select the best way of expressing a thought. The fourth grade test contained 14 items and 9 minutes were allowed for its completion; the seventh grade test contained 18 items and allowed 13 minutes for completion. The fourth grade was allowed 8 minutes for part C, written usage. The seventh grade test did not include written usage. Fourth graders were to recognize grammatical errors in 14 items. To recognize errors of punctuation and capitalization was the object of part D, for grade four and part C for grade seven. The fourth grade booklet had 12 items and allowed 8 minutes, and the seventh grade booklet had 20 items and allowed 11 minutes.

Mathematics

The mathematics test at both grade levels involved mathematical

reasoning, problem solving and computation. In addition, problems in the seventh grade test involved algebraic and geometric concepts. Pupils at each grade level had 30 minutes in which to answer 40 questions.

Composite Achievement

A composite achievement score was computed for each student. The composite score was obtained by averaging the individual's standard scores on the reading, the mechanics of written English, and the mathematics tests. The test scores were combined in this way so that each score would contribute equally to the average--despite the fact that the number of items was different on the three tests.

It should be noted that the word relationships test score was not included in the calculation of the composite achievement score. The word relationships are not a common subject of direct instruction and the word relationships score is believed to respond more slowly than the other scores to the influence of schooling. Therefore, it was excluded to focus the composite achievement score upon those aspects of basic skills achievement that respond most readily to change.

SECTION II

PRECAUTIONS IN THE INTERPRETATION OF PUPIL SCORES

A number of precautions should be observed in interpreting the individual pupil scores from the Michigan Educational Assessment Program. The following precautions are particularly important.

Measurement Error

All measurement is subject to error. Scores resulting from educational assessment tests are no exception. Therefore, in any use of the results an allowance for error should be made. The discussion on reliability in Section IV of this report contains estimates of the amount of error associated with scores on each of the assessment instruments on which individual pupil scores are reported.

The score of a pupil at any time will be the result of a number of influences. These influences include, but are not limited to: previous educational experience, effectiveness of the school program, teaching effectiveness, home environment, and peer culture. Furthermore, a pupil's score is likely to reflect the combined impact of these forces over many years--even at the fourth grade level.

Content is General

The educational assessment results provide a general idea of the basic skills achievement levels of each pupil compared to the basic skill levels of pupils throughout the state. In other words, the tests were purposely made general in their content so that they would be useful in the varied school situations that are to be found throughout the state. Thus, an effort has been made to focus upon the broader outcomes in reading, me-

chanics of writing, and mathematics sought by all schools. Because the tests are general, they will not reflect the skills and achievements that are taught less widely nor those unmeasurable in machine-scoreable tests.

Results Not Diagnostic

General achievement batteries are not designed to provide diagnostic information about individuals upon which specific instructional plans may be based. The Michigan educational assessment tests, being general achievement batteries, do not provide diagnostic information. Rather, they provide a general indication of a pupil's skill levels. If a particular student's results on the achievement tests suggest a problem, an appropriate diagnostic instrument may provide information useful for planning instruction to correct the problem. Information about diagnostic tests* in each skill area may be found in the series of Mental Measurements Yearbooks prepared by O.K. Buros. However, the administration of full diagnostic batteries to all pupils is generally unnecessary and prohibitively expensive.

Many areas of pupil achievement and development are not included in the educational assessment battery. It must be remembered that the goals of local schools and of the Michigan Department of Education are much broader than those assessed by this battery. Therefore, persons wishing to judge the overall achievement of pupils must look to additional measures for judgment in other areas of pupil development.

---

*A test is considered to be diagnostic if it provides a detailed analysis of strengths and weaknesses in an area. Such an analysis will often suggest possible causes for the deficiencies revealed and indicate possible remedial steps.

SECTION III

INTERPRETATION OF PUPIL SCORES

The purpose of this section is to provide the reader with information which will facilitate his interpretation of pupil scores. This section contains three parts: a description of the norming group, definitions of the terms necessary for understanding the pupil rosters and labels, and descriptions of the pupil rosters and labels.

Description of the Norming Group

The 1971-72 Michigan Assessment Battery was administered to public school students in the fourth and seventh grades. In non-graded programs pupils who were either identified as fourth or seventh graders or were in their fourth or seventh year beyond kindergarten were included. Also included were shared-time pupils who received instruction in the basic skills in the public school.

Excluded were all pupils receiving instruction in special classes for the handicapped in the content areas of reading, English, and mathematics; but remedial reading pupils were not excluded. Pupils receiving itinerant special education services (e.g., hard of hearing, physically handicapped, educable mentally retarded, emotionally disturbed) and who received instruction in the regular class program in the areas of reading, English and mathematics were included in the assessment.

Definition of Terms

This part defines the terms which are necessary for understanding the pupil rosters and labels. The defined terms are: mean, standard

deviation, standard scores, and percent below.

Mean

A mean score is an average of a set of scores. It is obtained by adding all of the scores and dividing the sum by the number of scores.

Standard Deviation

In addition to establishing a mean for a distribution of scores, it is often useful to know the "spread" of the scores. Two groups of scores could have the same mean but still be quite different. For example, one district might have children whose scores on composite achievement are very similar and have a mean score of fifty. In this district, the "spread" of scores would be small. Another district might have a number of children with high scores and a number of children with low scores and still have a mean score of fifty. In this district, however, the "spread" of scores would be large.

One common way of indicating the "spread" of a set of scores is to calculate the standard deviation. Usually about two-thirds of the scores will fall between one standard deviation above and one standard deviation below the mean. The larger the standard deviation, the larger will be the "spread" or variability among the scores of a distribution. In the example above, the district with similar scores would have a smaller standard deviation than would the district with the mixture of high and low scores.

Standard Scores

Standard scores are those that are derived from "raw" or response scores using the concepts of mean and standard deviation. In the Michigan Educational Assessment Program, standard scores were developed so that the pupil mean score from any assessment battery measure was fifty and the standard deviation was ten, when computed for all pupils at the same grade level. For example, a pupil with a standard score of forty on reading is one standard deviation below the state mean; a pupil with a score of sixty is one standard deviation above the mean; a pupil with a score of sixty-five is one and one-half standard deviations above the mean; and so forth.

Percent Below

A percent below corresponding to any given standard score is the percentage of pupils in the norm group who received lower scores. Thus, a percent below score of 75 on word relationships indicates that 75 percent of the pupils in the state received a lower score, and that 25 percent of the pupils in the state received the same or higher score; a percent below score of 16 would mean that 16 percent of the pupils in the state received a lower score and 84 percent of the pupils received the same or a higher score. And so on.

Description of Pupil Roster and Pupil Labels

The pupil roster and pupil labels yield exactly the same information in different formats. Therefore, only the pupil roster will be fully described.

ROSTER OF SCORES 1971-72

MICHIGAN DEPARTMENT OF EDUCATION - MICHIGAN EDUCATIONAL ASSESSMENT PROGRAM

SCHOOL NAME ABLE ELEMENTARY SCHOOL (3487) DISTRICT NAME MICHVILLE PUBLIC SCHOOLS (01-010) GRADE 4

STUDENT NAME			DATE OF BIRTH		SEX	WORD RELATIONSHIPS				READING				MECHANICS OF WRITTEN ENGLISH				MATH				COMPOSITE ACHIEVEMENT	
LAST	FIRST	M.I.	MONTH	YEAR		NUMBER RIGHT	NUMBER OF QUESTIONS	STANDARD SCORE	PERCENT BELOW	NUMBER RIGHT	NUMBER OF QUESTIONS	STANDARD SCORE	PERCENT BELOW	NUMBER RIGHT	NUMBER OF QUESTIONS	STANDARD SCORE	PERCENT BELOW	NUMBER RIGHT	NUMBER OF QUESTIONS	STANDARD SCORE	PERCENT BELOW	STANDARD SCORE	PERCENT BELOW
AABERSON	ALEX	M	07	62	M	25/45		49	46	30/50		49	41	35/55		54	60	20/40		46	33	50	46

Figure 1

Pupil Roster

Figure 1, on page 10, represents a pupil roster. An example from it will now be described. Alex M. Aaberson was born in July, 1962, is a male, and is in fourth grade. He received a "raw" score of 25 out of a possible 45 questions on the word relationships test. This "raw" score of 25 translates to a standard score of 49, and 46 percent of the pupils in the state received a lower score. He answered correctly 30 of 50 questions on the reading test for a standard score of 49, a score which exceeded 41 percent of the pupils in the state. On mechanics of written English, Alex got 35 of 55 questions correct for a standard score of 54, and 60 percent of the students in the state received a lower score. On mathematics, he got 20 of 40 questions correct for a standard score of 46, and 33 percent of the students in the state received a lower score. Finally his composite achievement standard score was 50, and 46 percent of the pupils in the state received a lower score. The remaining names and scores on the roster may be described in a similar manner. A pupil roster is provided for each school in the district which has a fourth and/or seventh grade.

Pupil Labels

The pupil label is displayed below and contains the same information as the pupil roster.

Figure 2

MICHIGAN EDUCATIONAL ASSESSMENT

NAME			GRADE	SEX	DATE OF BIRTH	
LAST	FIRST	M.I.			MONTH	YEAR
AABERSON	ALEX	M	4	M	07	62

1971-72

	WORD RELATIONSHIP	READING	MECHANICS OF WRITTEN ENG.	MATHEMATICS	COMPOSITE ACHIEVEMENT
NUMBER QUESTIONS RIGHT / IN TEST	25/45	30/50	35/55	20/40	
STANDARD SCORE	49	49	54	46	50
PERCENT BELOW	46	41	60	33	46

MICHIGAN EDUCATIONAL ASSESSMENT PROGRAM

SECTION IV

STATISTICAL CHARACTERISTICS OF THE 1971-72 MICHIGAN EDUCATIONAL ASSESSMENT BATTERY

The purpose of this section is to provide information concerning the statistical properties of the tests. The section is divided into three major parts. The first part will define the statistical terms which are necessary for the readers' understanding of test characteristics. The next part will discuss the statistical characteristics of the achievement measures at the fourth grade and seventh grade levels. The final part will evaluate the test characteristics and will indicate the areas in which additional information is needed.

Definition of Terms

This part defines the statistical terms: validity, reliability, difficulty, speededness, and standard error of measurement.

Validity

The validity of a test is an indication of the extent to which it measures what it is intended to measure. The most important type of validity for achievement tests is content validity. Content validity means that a test which claims to measure elementary mathematics, for example, should contain questions in mathematics and that those questions should be appropriate for the grade level for which the test is intended. The content validity of a test is dependent upon the extent to which the questions in the test constitute a representative sample of the

topics that comprise the subject tested.

It is virtually impossible to experimentally determine the content validity of a test of school achievement, or to report content validity as a numerical coefficient. The best available evidence of validity is found in the test itself and in a description of the procedures used to construct it.

Content validity is likely to be achieved if the development of the tests is the joint responsibility of specialists in test construction and specialists in the skills to be taught. The development of the achievement tests in the Michigan assessment battery has been described in Section I and could be expected to produce tests of high content validity.

Concurrent validity is another type of validity. Since it is reported as a correlation coefficient, it is useful in determining the extent to which two tests measure the same subject or characteristic.

Reliability

The reliability of a test provides an estimate of the test's consistency or stability. A test is reliable if it measures consistently whatever it measures; the most reliable tests yield relatively precise results for each student, and a student would receive roughly the same score on such a test if he could take it more than once under the same conditions. The reliability of a test is reported as a coefficient--that is, a two-place decimal figure. A reliability coefficient can range between .00 and 1.00. A test increases in reliability as the coefficient increases in value. An internal consistency method (Kuder-Richardson Formula #20) was used in estimating the reliability of the basic skills tests in the Michigan educational assessment battery. This method employs

information about the length of the test and the extent to which the questions in the test contribute mutually confirming or consistent information.

Difficulty

The difficulty of a test is an indication of how well suited it is to the ability of the group being tested. A test of middle difficulty is appropriate when the group being tested is heterogeneous, such as the group being tested in the Michigan Educational Assessment Program. A test is of middle difficulty when the mean score is near the point midway between a perfect score and the expected chance score. If the test includes 50 items a perfect score would, of course, be 50. The expected chance score, when the 50 items each offer four alternative answers, is 1/4 of 50 or 12.5. Hence the ideal mean (midway between chance and perfect) is 31.25. Expressed as a percentage, this would be 62.5%. In the Michigan educational assessment battery, a percentage significantly higher than 62.5% would indicate an easy test; a percentage significantly lower than 62.5% would indicate a difficult one.

Speededness

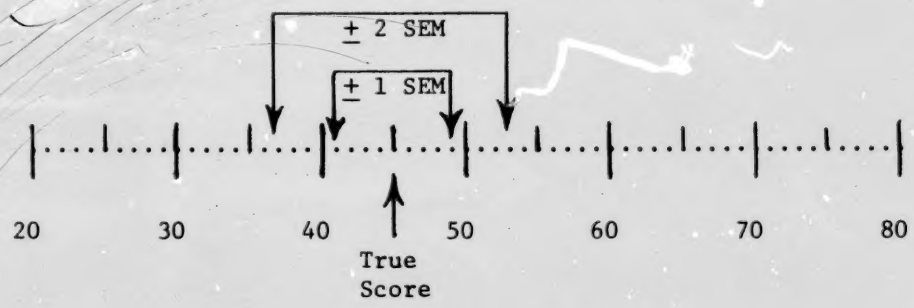
Speededness is a measure of the extent to which test performance is affected by the time limit placed on the test. The criterion used in judging the speededness of a test is two-fold: the proportion of students who answered (1) the last question and (2) the question that is three-quarters of the way through the test. This does not mean that the students answered all the questions up to these points; it means that they reached these particular items. If 80 percent of the students complete the last item and virtually all the students complete 75 percent

of the items, the test is usually judged to be unspeeded. It should be noted that it is desirable for tests in the Michigan Educational Assessment Program to be unspeeded.

Standard Error of Measurement

The standard error of measurement yields an index of the error associated with a test score. It may be used to form an error band around a score extending from a given number of standard errors of measurement above the score to the same number of standard errors of measurement below the score. Figure 3 illustrates the error band around the true score for a standard error of measurement of 4 score units.

Figure 3



If many individuals with the same true score of 45 were tested about 68 percent of them would receive actual scores not more than one standard error of measurement (SEM) from 45, and about 95 percent would receive actual scores not more than two standard errors of measurement (SEM) from 45.

Statistical Characteristics of the Fourth Grade Basic Skills Battery

A representative sample of 1,000 student tests was drawn from the number of fourth grade students who participated in the 1971-72 Michigan Educational Assessment Program. This sample's responses were used to analyze the characteristics of the fourth grade educational assessment battery.

The Test Analysis Sample

The means and standard deviations for the sample and for the entire population are shown in Table 1 along with the number of questions contained in each test. From the table it can be seen that the characteristics of the sample conform closely to the characteristics of the entire fourth grade population. Therefore the results presented in the following tables may be generalized to the fourth grade population. Detailed information concerning the distributions of raw and standard scores is reported in the Appendix.

Table 1

STATISTICS ON THE GRADE FOUR 1971-72 EDUCATIONAL ASSESSMENT TESTS FOR THE TEST ANALYSIS SAMPLE AND FOR ALL FOURTH GRADERS TESTED

		Sample		Population	
Score	Number of Questions	Mean	Standard Deviation	Mean	Standard Deviation
Word Relationships	45	25.2	9.2	25.7	9.3
Reading	50	30.7	10.9	31.0	10.9
Mechanics of Written English	55	30.9	10.8	31.1	10.7
Mathematics	40	22.9	7.4	23.0	7.3

Reliability

Table 2 presents the reliability coefficients and standard errors of measurement. The reliability coefficients were: .906 for word relationships, .931 for reading, .916 for mechanics of written English, .872 for mathematics and .964 for composite achievement. The standard errors of measurement expressed in raw score units were: 2.8 on word relationships, 2.9 on reading, 3.1 on mechanics of written English, 2.6 on mathematics, and (not defined) on composite achievement. The standard errors of measurement expressed in standard score units were: 3.0 on word relationships, 2.6 on reading, 3.0 on mechanics of written English, 3.6 on mathematics, and 1.8 on composite achievement. The reliability coefficients are adequately high and the standard errors of measurement are adequately low to justify reporting of individual student scores.

Table 2

RELIABILITY COEFFICIENTS AND STANDARD ERRORS OF MEASUREMENT FOR A SAMPLE OF 1,000 FOURTH GRADERS ON THE 1971-72 EDUCATIONAL ASSESSMENT BATTERY

Score	Reliability Coefficient	Standard Error of Measurement	
		Raw Score Units	Standard Score Units
Word Relationships	.906	2.8	3.0
Reading	.931	2.9	2.6
Mechanics of Written English	.916	3.1	3.0
Mathematics	.872	2.6	3.6
Composite Achievement	.964	(not defined)	1.8

Difficulty and Speededness

Table 3 presents the difficulty and speededness indices. The difficulty indices on word relationships, reading, mechanics of written English, and mathematics are within an acceptable range of middle difficulty. The table does indicate that the word relationships test is speeded.

Table 3

DIFFICULTY AND SPEEDEDNESS FOR A SAMPLE OF 1,000 FOURTH GRADERS ON THE 1971-72 EDUCATIONAL ASSESSMENT BATTERY

Score	Average Item Difficulty (Percent passing)	Speededness: Percent Reaching Last Item	Speededness: Percent Completing 75% of test
Word Relationships	57.1	68.6	87.2
Reading	62.0	82.1	95.8
Mechanics of Written English	56.5	*	*
a. Spelling	n.a.	90.0	95.6
b. Effectiveness of Written Expression	n.a.	86.2	95.0
c. Written Usage	n.a.	91.2	95.6
d. Punctuation and Capitalization	n.a.	91.4	95.8
Mathematics	57.5	85.9	96.4

n.a. - not available
* - not applicable

Statistical Characteristics of the Seventh Grade Basic Skills Battery

A representative sample of 1,000 student tests was drawn from the number of seventh grade students who participated in the 1971-72 Michigan Educational Assessment Program. This sample's responses were used to analyze the characteristics of the seventh grade educational assessment battery.

The Test Analysis Sample

The means and standard deviations for the sample and for the entire population are shown in Table 4 along with the number of questions in each test. From the table it can be seen that the characteristics of the sample conform closely to the characteristics of the entire seventh grade population. Therefore the results presented in the following tables may be generalized to the seventh grade population. Detailed information concerning the distribution of raw and standard scores is reported in the Appendix.

Table 4

STATISTICS ON THE GRADE SEVEN 1971-72 EDUCATIONAL ASSESSMENT TESTS FOR THE TEST ANALYSIS SAMPLE AND FOR ALL SEVENTH GRADERS TESTED

		Sample		Population	
Score	Number of Questions	Mean	Standard Deviation	Mean	Standard Deviation
Word Relationships	50	29.5	8.7	29.3	8.8
Reading	50	34.2	10.0	33.9	10.0
Mechanics of Written English	60	35.4	10.3	35.1	10.5
Mathematics	40	23.9	8.0	23.8	8.1

Reliability

Table 5 presents the reliability coefficients and standard errors of measurement. The reliability coefficients were: .887 for word relationships, .921 for reading, .900 for mechanics of written English, .892 for mathematics and .962 for composite achievement. The standard errors of measurement expressed in raw score units were: 2.9 on word relationships, 2.8 on reading, 3.3 on mechanics of written English, 2.6 on mathematics, and (not defined) on composite achievement. The standard errors of measurement expressed in standard score units were: 3.4 on word relationships, 2.8 on reading, 3.1 on mechanics of written English, 3.3 on mathematics, and 1.8 on composite achievement. The reliability coefficients are adequately high and the standard errors of measurement are adequately low to justify reporting of individual student score.

Table 5

RELIABILITY COEFFICIENTS AND STANDARD ERRORS OF MEASUREMENT FOR A SAMPLE OF 1,000 SEVENTH GRADERS ON THE 1971-72 EDUCATIONAL ASSESSMENT BATTERY

		Standard Error of Measurement	
Score	Reliability Coefficient	Raw Score Units	Standard Score Units
Word Relationships	.887	2.9	3.4
Reading	.921	2.8	2.8
Mechanics of Written English	.900	3.3	3.1
Mathematics	.892	2.6	3.3
Composite Achievement	.962	(not defined)	1.8

Difficulty and Speededness

Table 6 presents the difficulty and speededness indices. The difficulty indices on word relationships, reading, mechanics of written English, and mathematics are within an acceptable range of middle difficulty. All of the tests are unspeeded.

Table 6

DIFFICULTY AND SPEEDEDNESS FOR A SAMPLE OF 1,000 SEVENTH GRADERS ON THE 1971-72 EDUCATIONAL ASSESSMENT BATTERY

Score	Average Item Difficulty (Percent passing)	Speededness: Percent Reaching Last Item	Speededness: Percent Completing 75% of Test
Word Relationships	58.6	82.4	96.0
Reading	67.8	93.9	98.9
Mechanics of Written English	58.5	*	*
a. Spelling	n.a.	94.2	98.8
b. Effectiveness of Written Expression	n.a.	96.3	98.7
c. Punctuation and Capitalization	n.a.	94.9	97.8
Mathematics	59.5	89.5	96.5

n.a. - not available
* - not applicable

Summary

A perusal of tables 1 through 6 indicates the statistical strength of the achievement batteries. However, certain types of statistical data are presently not available in the Michigan Educational Assessment Program.

The Michigan Educational Assessment Program has provided reasonable content validity in the achievement battery. Programs are currently underway to make the content of the assessment battery even more relevant to Michigan's curricula. Concurrent validity information is available upon request.

APPENDIX

Table of Word Relationships Raw Scores, Standard Scores and Percents Below

	Grade 4			Grade 7		
Raw Score	Standard Score	Number of Pupils	Percent Below	Standard Score	Number of Pupils	Percent Below
50				74	20	100.0
49				72	106	99.9
48				71	346	99.7
47				70	670	99.3
46				69	1150	98.6
45	71	307	99.8	68	1670	97.6
44	70	889	99.3	67	2283	96.2
43	69	1722	98.2	66	2878	94.4
42	67	2347	96.7	64	3645	92.1
41	66	3076	94.8	63	4092	89.6
40	65	3576	92.6	62	4724	86.7
39	64	3878	90.1	61	5105	83.6
38	63	4261	87.5	60	5645	80.1
37	62	4458	84.7	59	6019	76.4
36	61	4553	81.9	58	6266	72.5
35	60	4729	78.9	56	6285	68.6
34	59	4849	75.9	55	6451	64.7
33	58	4859	72.8	54	6591	60.6
32	57	5088	69.7	53	6496	56.6
31	56	5061	66.5	52	6407	52.6
30	55	5158	63.3	51	6358	48.7
29	54	5205	60.0	50	6225	44.9
28	52	5347	56.7	48	6114	41.1
27	51	5600	53.2	47	5859	37.5
26	50	5799	49.6	46	5741	34.0
25	49	5835	46.0	45	5412	30.6
24	48	5861	42.3	44	5236	27.4
23	47	5916	38.6	43	5051	24.3
22	46	5824	35.0	42	4932	21.2
21	45	5622	31.5	41	4895	18.2
20	44	5656	27.9	39	4600	15.4
19	43	5445	24.5	38	4312	12.7
18	42	5229	21.3	37	3879	10.3
17	41	4854	18.2	36	3455	8.2
16	40	4358	15.5	35	3008	6.4
15	39	4125	12.9	34	2463	4.8
14	37	3884	10.5	33	2010	3.6
13	36	3357	8.4	31	1617	2.6
12	35	3011	6.5	30	1221	1.9
11	34	2548	5.0	29	914	1.3
10	33	2058	3.7	28	690	0.9
9	32	1659	2.6	27	500	0.6
8	31	1309	1.8	26	322	0.4
7	30	956	1.2	25	207	0.2
6	29	687	0.8	23	136	0.2
5	28	510	0.5	22	101	0.1
4	27	349	0.3	21	62	0.1
3	26	217	0.1	20	39	0.0
2	25	127	0.0	19	28	0.0
1	24	56	0.0	18	11	0.0
0	22	19	0.0	17	7	0.0

Table of Reading Raw Scores, Standard Scores and Percents Below

	Grade 4			Grade 7		
Raw Score	Standard Score	Number of Pupils	Percent Below	Standard Score	Number of Pupils	Percent Below
50	67	214	99.9	66	1496	99.1
49	67	906	99.3	65	3138	97.1
48	66	1867	98.1	64	4296	94.5
47	65	3003	96.3	63	5251	91.3
46	64	3988	93.8	62	5661	87.8
45	63	4750	90.8	61	5920	84.1
44	62	5370	87.5	60	6308	80.2
43	61	5818	83.8	59	6421	76.3
42	60	5953	80.1	58	6361	72.4
41	59	6072	76.3	57	6386	68.4
40	58	6046	72.5	56	6096	64.7
39	57	5880	68.9	55	6061	60.9
38	56	5816	65.2	54	5890	57.3
37	55	5580	61.8	53	5755	53.8
36	55	5469	58.4	52	5490	50.4
35	54	5150	55.1	51	5314	47.1
34	53	4961	52.0	50	5159	43.9
33	52	4778	49.1	49	4973	40.9
32	51	4585	46.2	48	4737	37.9
31	50	4388	43.5	47	4663	35.1
30	49	4041	40.9	46	4404	32.3
29	48	4089	38.4	45	4294	29.7
28	47	3885	36.0	44	4116	27.2
27	46	3515	33.8	43	3959	24.7
26	45	3452	31.6	42	3718	22.4
25	44	3422	29.5	41	3631	20.2
24	44	3239	27.5	40	3499	18.0
23	43	3224	25.4	39	3198	16.1
22	42	3216	23.4	38	3079	14.2
21	41	3159	21.5	37	2970	12.3
20	40	3139	19.5	36	2825	10.6
19	39	3187	17.5	35	2641	9.0
18	38	3277	15.5	34	2465	7.5
17	37	3351	13.4	33	2215	6.1
16	36	3337	11.3	32	2041	4.8
15	35	3242	9.3	31	1805	3.7
14	34	3093	7.3	30	1623	2.7
13	33	2818	5.6	29	1277	1.9
12	33	2481	4.0	28	1038	1.3
11	32	1922	2.8	27	776	0.8
10	31	1493	1.9	26	510	0.5
9	30	1054	1.3	25	302	0.3
8	29	725	0.8	24	208	0.2
7	28	479	0.5	23	143	0.1
6	27	283	0.3	22	65	0.1
5	26	213	0.2	21	37	0.0
4	25	130	0.1	20	29	0.0
3	24	84	0.1	19	12	0.0
2	23	54	0.0	18	13	0.0
1	22	29	0.0	17	4	0.0
0	21	8	0.0	16	2	0.0

Table of Mechanics of Written English Raw Scores, Standard Scores and Percents Below

	Grade 4			Grade 7		
Raw Scores	Standard Score	Number of Pupils	Percent Below	Standard Score	Number of Pupils	Percent Below
60				74	39	100.0
59				73	122	99.9
58				72	282	99.7
57				71	449	99.5
56				70	681	99.0
55	72	73	100.0	69	976	98.4
54	72	246	99.8	68	1275	97.6
53	71	519	99.5	67	1713	96.6
52	70	817	99.0	66	2023	95.3
51	69	1271	98.2	65	2446	93.8
50	68	1495	97.2	64	2789	92.1
49	67	2016	96.0	63	3205	90.1
48	66	2391	94.5	62	3711	87.9
47	65	2823	92.7	61	4140	85.3
46	64	3289	90.7	60	4436	82.6
45	63	3460	88.5	59	4638	79.7
44	62	3832	86.1	58	4966	76.7
43	61	4129	83.5	58	5274	73.4
42	60	4321	80.8	57	5560	70.0
41	59	4628	77.9	56	5669	66.5
40	58	4690	75.0	55	5835	62.9
39	57	4830	72.0	54	5755	59.3
38	57	4782	69.0	53	5790	55.8
37	56	4825	66.0	52	5700	52.3
36	55	5182	62.8	51	5704	48.8
35	54	4975	59.7	50	5544	45.3
34	53	5023	56.5	49	5281	42.1
33	52	5009	53.4	48	5105	38.9
32	51	4884	50.3	47	4935	35.9
31	50	4949	47.2	46	4819	32.9
30	49	4682	44.3	45	4612	30.1
29	48	4895	41.3	44	4176	27.5
28	47	4755	38.3	43	4167	24.9
27	46	4550	35.5	42	3980	22.5
26	45	4669	32.5	41	3746	20.2
25	44	4540	29.7	40	3502	18.0
24	43	4452	26.9	39	3355	16.0
23	42	4305	24.2	39	3177	14.0
22	41	4223	21.6	38	2940	12.2
21	41	4098	19.0	37	2759	10.5
20	40	3964	16.6	36	2608	8.9
19	39	3852	14.2	35	2573	7.3
18	38	3576	11.9	34	2184	6.0
17	37	3255	9.9	33	1971	4.7
16	36	3113	7.9	32	1757	3.7
15	35	2688	6.3	31	1508	2.7
14	34	2340	4.8	30	1188	2.0
13	33	1982	3.6	29	937	1.4
12	32	1556	2.6	28	765	1.0
11	31	1248	1.8	27	511	0.6
10	30	960	1.2	26	336	0.4
9	29	649	0.8	25	244	0.3
8	28	418	0.5	24	148	0.2
7	27	323	0.3	23	104	0.1
6	26	203	0.2	22	68	0.1
5	26	132	0.1	21	43	0.1
4	25	78	0.1	21	31	0.0
3	24	68	0.0	20	21	0.0
2	23	32	0.0	19	23	0.0
1	22	33	0.0	18	9	0.0
0	21	7	0.0	17	5	0.0

Table of Mathematics Raw Scores, Standard Scores and Percents Below

	Grade 4			Grade 7		
Raw Score	Standard Score	Number of Pupils	Percent Below	Standard Score	Number of Pupils	Percent Below
40	73	94	99.9	70	287	99.8
39	72	313	99.7	69	1155	99.1
38	71	726	99.3	68	2301	97.7
37	69	1231	98.5	66	3434	95.6
36	68	1959	97.3	65	4523	92.8
35	67	2840	95.5	64	5119	89.6
34	65	3754	93.2	63	5593	86.2
33	64	4605	90.3	61	5691	82.6
32	62	5360	86.9	60	5923	79.0
31	61	5953	83.2	59	6155	75.2
30	60	6667	79.0	58	6153	71.4
29	58	7020	74.7	56	6189	67.6
28	57	7348	70.1	55	6251	63.7
27	56	7542	65.3	54	6272	59.8
26	54	7561	60.6	53	6318	55.9
25	53	7729	55.8	51	6332	52.0
24	51	7489	51.1	50	6410	48.1
23	50	7455	46.4	49	6150	44.3
22	49	7318	41.9	48	6161	40.5
21	47	7203	37.3	47	6180	36.7
20	46	6983	33.0	45	6172	32.8
19	45	6708	28.8	44	6086	29.1
18	43	6418	24.8	43	6010	25.4
17	42	5977	21.0	42	5760	21.8
16	40	5721	17.5	40	5619	18.4
15	39	5161	14.2	39	5317	15.1
14	38	4768	11.2	38	5016	12.0
13	36	4141	8.7	37	4527	9.2
12	35	3536	6.4	35	3994	6.7
11	34	2924	4.6	34	3384	4.6
10	32	2396	3.1	33	2659	3.0
9	31	1757	2.0	32	1935	1.8
8	29	1229	1.2	30	1258	1.0
7	28	775	0.8	29	754	0.5
6	27	534	0.4	28	439	0.3
5	25	297	0.2	27	232	0.1
4	24	193	0.1	25	107	0.1
3	23	84	0.1	24	61	0.0
2	21	58	0.0	23	20	0.0
1	20	34	0.0	22	17	0.0
0	18	14	0.0	20	5	0.0

Table of Composite Achievement Standard Scores and Percents Below

Standard Score	Grade 4 Number of Pupils	Grade 4 Percent Below	Grade 7 Number of Pupils	Grade 7 Percent Below
71	7	100.0	0	100.0
70	92	99.9	9	100.0
69	302	99.7	116	99.9
68	654	99.3	445	99.6
67	1170	98.6	943	99.1
66	1735	97.5	1651	98.0
65	2411	96.0	2375	96.6
64	3138	94.0	3036	94.7
63	3864	91.6	3762	92.3
62	4247	88.9	4454	89.6
61	4903	85.8	5129	86.4
60	5432	82.4	5458	83.0
59	5685	78.8	5792	79.4
58	5948	75.1	5858	75.7
57	6043	71.3	6136	71.9
56	6158	67.4	6319	68.0
55	6227	63.5	6430	64.0
54	6184	59.6	6285	60.1
53	5895	55.9	6158	56.3
52	5573	52.4	6045	52.5
51	5411	49.0	5987	48.8
50	5401	45.6	5829	45.2
49	5299	42.3	5589	41.7
48	4965	39.2	5440	38.3
47	4923	36.1	5214	35.1
46	4887	33.0	5085	31.9
45	4791	30.0	4775	28.9
44	4558	27.1	4648	26.0
43	4452	24.3	4548	23.2
42	4211	21.7	4275	20.6
41	4331	18.9	4130	18.0
40	4217	16.3	3933	15.6
39	4055	13.7	3728	13.2
38	3842	11.3	3503	11.1
37	3677	9.0	3150	9.1
36	3351	6.9	2995	7.2
35	3006	5.0	2654	5.6
34	2448	3.5	2332	4.1
33	1942	2.2	2051	2.9
32	1433	1.3	1600	1.9
31	911	0.8	1202	1.1
30	508	0.5	838	0.6
29	279	0.3	457	0.3
28	180	0.2	273	0.1
27	92	0.1	126	0.1
26	87	0.1	50	0.0
25	44	0.0	34	0.0
24	26	0.0	14	0.0
23	6	0.0	11	0.0
22	4	0.0	2	0.0
21	1	0.0	2	0.0

Raw and Standard Score Means and Standard Deviations
For the Standardization Group

Test	Statistic	Grade 4 Raw Score	Grade 4 Standard Score	Grade 7 Raw Score	Grade 7 Standard Score
Word Relationships	Mean	25.7	50.0	29.3	50.0
	Standard Deviation	9.3	10.0	8.8	10.0
Reading	Mean	31.0	50.0	33.9	49.9
	Standard Deviation	10.9	10.0	10.0	10.0
Mechanics of Written English	Mean	31.1	50.0	35.1	50.0
	Standard Deviation	10.7	10.0	10.5	10.0
Mathematics	Mean	23.0	50.0	23.8	50.0
	Standard Deviation	7.3	10.0	8.1	10.0
Composite Achievement	Mean	*	50.0	*	50.0
	Standard Deviation	*	9.3	*	9.1

* - not applicable